Response to referee comments about "Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC example" by Robert J. Andres, Thomas A. Boden, and David M. Higdon.

The authors thank Ray Nassar and an anonymous referee for their thoughtful comments on the manuscript. Following are their review comments (in black text) and our responses (in blue text). The authors have made many changes to the manuscript originally submitted to ACPD in response to these comments.

Ray Nassar comments

Andres, Boden and Higdon "Gridded uncertainty in fossil fuel carbon dioxide emission maps, a CDIAC example" presents a new method to derive gridded uncertainty in CDIAC fossil fuel CO_2 emission maps. The scientific significance of this work is high, since it is useful and important work for carbon cycle science and perhaps also climate policy. The scientific quality is high with the present work pioneering new methods for this field. The presentation quality is good, but there is room for some improvement here.

The comments of this reviewer and the next have improved the presentation quality if the ACP-destined manuscript. The ACP-destined text has been modified.

Quantified uncertainty in fossil fuel CO_2 emission data is something that has been needed by the carbon cycle science community for some time. Andres et al. (2014) took a first step in deriving a global uncertainty value, but this manuscript goes much further with actual gridded uncertainty maps that users of the gridded emission maps will soon be able to access. As CO_2 modellers use the gridded uncertainties, this manuscript will serve as a valuable reference to understand the details of how the uncertainties were derived and any limitations that they have. The manuscript also presents a revision to their gridded emission approach, specifically the implementation of time-varying population distributions for 1990-2011. One important issue that the author's did not address in enough detail was the reliability of population as a proxy for fossil fuel (and cement) CO_2 emissions. I have some suggestions for improvements to that discussion, but a more thorough analysis may perhaps be best addressed in a future paper. Overall, I would be supportive of publication of this manuscript in Atmospheric Chemistry and Physics, provided that some revisions are made according to the suggestions that follow.

The authors agree with the comments of the reviewer. The population proxy has been discussed elsewhere and within this manuscript. It is not an ideal proxy. All proxies have their limitations. The authors believe what is really needed are rigorous uncertainty assessments of the various FFCO2 mapping approaches. Then, the strengths and weakness of the various approaches can be better understood and, maybe, remedied. The manuscript has not been modified in response to this comment.

Lines 42-50: The authors' discussion of the reduction in uncertainty for other components of the carbon cycle might leave the impression on a reader that the uncertainties on these components is smaller than it actually is. This is especially true for the terrestrial biospheric CO_2 flux, which is usually calculated as the residual of all other fluxes including fossil fuel (and cement) CO_2 emissions. This fact should be mentioned since in effect, good biospheric flux estimates (and knowledge of their uncertainties) depend on good FFCO2 estimates (and knowledge of their uncertainties). Secondly, I am surprised that the authors fail to mention in the introduction the fact that the uncertainty in FFCO2 may actually be growing for two reasons: 1) the overall magnitude of FFCO2 emissions has continued to grow (about 2.5 PgC/yr in 1960 to about 10 PgC/yr now) as shown in Figure 10 of Andres et al. (2014), so any fixed percent uncertainty would translate to a larger absolute value, and 2) the fact the growth has mainly occurred in countries with higher FFCO2 uncertainties, according to Table 2 in Andres et al. (2014). I would recommend adding these points to the introduction (in the authors' own words).

Two sentences added to first paragraph of Introduction to address role of uncertainty in component fluxes of global carbon cycle and effect of FFCO2 fluxes on terrestrial biosphere fluxes. One sentence added regarding the two reasons the reviewer noted. The ACP-destined text has been modified.

Lines 95-101: I find the wording "stock maps" to be odd, definitely not the standard in the field. If they mean maps of "atmospheric CO_2 concentrations", it would be better to simply say that (or to be more exact "atmospheric CO_2 mixing ratios" or "atmospheric CO_2 column-averaged dry air mole fractions").

The word “stock” removed. The ACP-destined text has been modified.

Lines 199-201: The authors should add once sentence to elaborate on what they mean by nighttime and daytime populations to assist readers. I suspected that I knew what they meant, but had to look up the terms to verify.

The ACP-destined text has been modified.

Lines 309-313. I do not think that a sub-national border is relevant here since the emission maps do not account for tabular data from individual sub-national units (states/provinces etc.). Is that the only example of this issue? This example may be relevant to mention with respect to lines 381-384.

This is not the only example of the surveying error issue. More examples were not given due to text length. The issue largely arises over time when historical physical descriptions of borders succumb to geological processes (e.g., rivers change course) or technological improvements (e.g., surveying techniques). The example is kept as this will be an issue for finer spatial scale maps as the text already mentions here and later in lines 381-384 that the reviewer mentions.

The manuscript has not been modified in response to this comment.

Line 420: I recommend changing "Vulcan data product" to "Vulcan data product for the United States" to emphasize that this was not a global check.

The ACP-destined text has been modified.

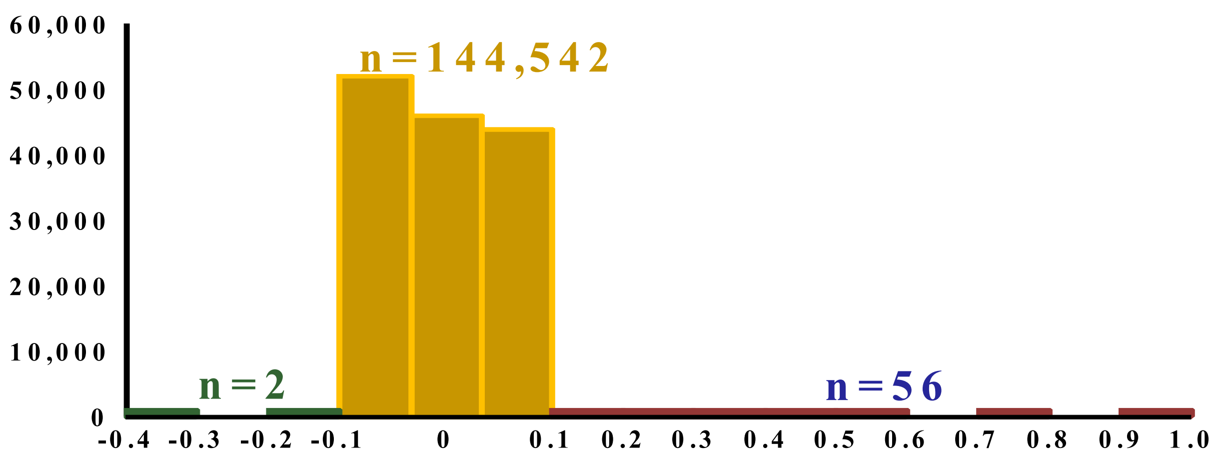
Lines 458-460: For larger grid box sizes, it would be possible to test this by aggregating the gridded 1x1 data, for example to 2x2 or 5x5, and then repeating the comparison. They authors may want to try this. I would expect such a test to confirm that the population-FFCO2 emission relationship is more reliable for coarser spatial resolution. However, since the relationship is very likely non-linear, I would not expect this approach to be able to provide any quantitative information about the population-FFCO2 relationship at the finer spatial resolutions (0.25, 0.1 and 0.01 degrees).

This comment is similar to the L735 comment of the anonymous reviewer. A simple aggregation of the present 1x1 gridded data is not representative of the CDIAC mapping process (as it would mix emissions from different countries which were originally mapped with different national per capita emissions). A more complete treatment would also entail modifying the geography map so that it agrees with the population map. This would then entail modifying the carbon distribution algorithm for handling the new spatial resolutions. The new spatial resolutions would be particularly difficult for small countries. The calculations necessary to test the comment of the reviewer have not been performed as 1) it is outside the scope of this manuscript, 2) if done, the results are not representative of an actual data product available, and 3) many decisions would be necessary along the calculation pathway that would need to be justified and explained for the calculation results to have merit. Different decisions could lead to very different results.

The authors agree that the population-FFCO2 relationship becomes more problematic at spatial resolutions finer the one degree (e.g., (0.25, 0.1 and 0.01 degrees).

The manuscript has not been modified in response to this comment.

Lines 497-506: I am a bit surprised by the conclusion that year of the population map did not matter much. Using the United States as an example, an old map might not have the correct East-West weighting since cities in the western part of the country have had more recent growth increases. In the developing world, there is a trend toward urbanization, with formal rural agriculturally-based populations resettling in cities. I would like to see at least one figure somehow quantifying the differences from using temporally-dependent population maps, so that the reader can judge the scale of the differences. Perhaps the difference were not large since the time period 1990-2012 was limited rather than spanning the whole data set period.



An earlier version of the manuscript had the above figure, but it was deleted for the ACPD submission as the text sufficiently describes the results.

The y-axis is number of grid cells and the x-axis is absolute change between successive years in within country relative population density (in percent, ten years of data, 1990-2000). The conclusion is most (>99%) populated grid cells show less than a 1% change in relative population density over a 10 year period. The scope of the analysis is global and thus incorporates a wide variety of urbanization trends over the decade. The reviewer is correct that a longer time series analysis may reveal larger differences, but the authors do not have annually varying population data prior to 1990.

The manuscript has not been modified in response to this comment.

Line 552: "borders in water-dominated areas" should be "borders and in water-dominated areas"

Geography does not dominate land-land borders (e.g., along the 49the parallel between the U.S. and Canada), but along some land-water borders (e.g., islands and coastlines). The ACP-destined text has been modified to clarify this.

There are multiple places in the manuscript (in addition to lines 458-460), where the issue of the population-FFCO2 emission relationship, the validity or per capita emissions, or the use of population as a proxy for FFCO2 emissions (versus other proxy data) is raised. The neglect of point sources in the current method was also mentioned with reference to the work of Singer et al. (2014). The spatial distributions from distributing emissions with population as a proxy and not accounting for large point sources could possibly be the largest uncertainty in the CDIAC 1x1 emission dataset, but this uncertainty is also the most difficult to address since the "true" FFCO2 emission values at fine spatial scales are simply not known. Nassar et al. (2013, JGR, 118, 917-933) compared the CDIAC 1x1 data (which uses a population-based proxy) and the ODIAC 1x1 data (which uses large point sources and a nighttime lights based proxy) over Canada against tabular reported activity-based FFCO2 at the provincial scale. The sum of the gridded data in each of 10 Canadian provinces and the combined northern territories was assigned a scale factor for each of dataset (CDIAC or ODIAC) that would give agreement with the tabular reported provincial total (see Figure 5 and Table 4). This test showed that less scaling was required (scale factors were usually closer to unity) when point sources and nightlights were used compared with sing population. It is expected that this result would hold for other large countries (like the US or China) where there are regional differences in methods of energy generation including non-CO_2 emitting methods like nuclear or renewable energy (hydroelectricity, wind, solar, geothermal, tidal etc.) or there are large variations in climate, transportation systems or building style within a country (e.g. New York vs. Texas) resulting in large differences in per capita energy consumption. While the analysis in Nassar et al. (2013) is not a complete analysis or comparison of all proxy methods globally, a summary of the above finding and citation of the paper is surely warranted in this manuscript.

The authors are certainly aware of the Nassar et al (2013) paper as one of the authors of this manuscript is also a coauthor on the Nassar et al. paper. However, the reviewer’s comment may give a misleading summary of the quality of the ODIAC methodology. Looking at Nassar et al.

(2013) Table 4, of the 11 sub-national units listed, six units have a ratio closer to one in the ODIAC methodology and five units have a ratio closer to one in the CDIAC methodology. For the national unit, Canada, the CDIAC methodology has a ratio closer to one than the ODIAC methodology. This is not to say that the CDIAC methodology is superior or inferior to the ODIAC methodology, it is a simple accounting of the ratios closest to unity (where unity equals agreement with the Canadian national inventory). What this accounting ignores is the within-political-unit distribution. One hopes that the large point source methodology used in ODIAC gives a superior emission distribution than the CDIAC population-only distribution, but there is not definitive proof of such superiority.

As stated in the fourth paragraph of the Introduction, this ACPD manuscript does not compare the CDIAC maps with maps produced by other authors. The points raised by the reviewer in connection with the Nassar et al. (2013) paper is essentially such a comparison. Nassar et al. (2013) state "suggests that this approach [i.e., ODIAC] is an improvement relative to CDIAC" (p. 925). This ACPD manuscript attempts to go beyond the side-by-side pictures of Nassar et al (2013) and provide a statistical basis (via calculated uncertainties) to compare different maps. Once such uncertainties are calculated for other maps, then a more meaningful assessment of various map methodologies can be accomplished.

The authors believe the ACPD manuscript is quite clear in 1) the CDIAC approach is a population proxy approach, 2) other maps include other proxies, and 3) robust estimates of uncertainty are necessary to rigorously compare maps produced by different methodologies. As these three points are already in the ACPD manuscript and Nassar et al. (2013) does not incorporate point # 3, the authors exclude referring to Nassar et al. (2013) in reference to this comment, but do include a call to the paper elsewhere in the ACP-destined manuscript in support of another issue.

While not explicitly including the requested reference to the Nassar et al. (2013) paper, the ACP-destined text has been modified in response to this comment by pointing out the limitations of the Fig. 8 results.

Figure 3. According to the text (line 267-270), the monthly uncertainty is a constant at 12.8

Yes, and the main effect of this shows up in Fig. 3 and its derivative products (e.g., Table 3,). The manuscript has not been modified in response to this comment.

Figures 3,7,9, and 10. The color-bars all contain two very similar shades of blue that are very difficult to distinguish. The color-bar for Figure 11 nicely avoids this problem. Can the authors revise their colour-bars to avoid colours that are too similar as they have nicely done in Figure 11.

Figures 3, 7, 9, and 10 colors bars changed to remove similar blues. The ACP-destined text has been modified.

Figure 8. Units for FFCO2 should be given. Removing some zeros in the FFCO2 or populations would also enhance clarity.

The ACP-destined text has been modified.

Figure 12. The label "Urban - vertical extent?" is confusing and not discussed anywhere. Either explain or just replace with "Urban".

In the main body of the text, two sentences after the sentence with the first call to Fig. 12, is the text "Much effort is now being directed to produce urban scale maps, their uncertainty at present is largely unknown" which is the discussion of the vertical extent comment. Text has been added to the Fig. 12 caption. The ACP-destined text has been modified.

Anonymous referee comments

General Comments:

The subject of this paper is very important for understanding carbon balance at sub-national scales, because the uncertainty of fossil fuel-CO2 emissions (FF) is a large (if not the largest) component of the uncertainty budget at annual scales. Thus, trying to quantify annual net ecosystem exchange is linked closely to FF and their uncertainty. Moreover, FF and their uncertainty are critically important in their own right from a policy/mitigation perspective.

We agree with the reviewer that at annual and global scales, FFCO2 uncertainty is large (either in isolation or in comparison with other major components of the global carbon cycle). The manuscript has not been modified in response to this comment.

The paper is primarily a methodological description of the gridded uncertainty calculation. However, for a methodological paper, I found the calculation details hard to follow. This was especially true in section 4.3, which describes the calculation of uncertainty related to distributing national-level FF by use of a population density proxy. This is the heart of the paper, and needs to be described more clearly. In fact, at this point I can't be sure if the results are justified, because I don't understand exactly how the uncertainty was calculated.

The reviewer offered below many constructive comments to clarify section 4.3 and elsewhere. Most of these reviewer comments have been incorporated into the ACP-destined text; explanations are given for comments not incorporated. The ACP-destined text has been modified.

My confusion with the detailed methods notwithstanding, I have a concern with the general approach. The main question one wants to answer when using population as a within-country proxy for FF is: to what extent are emissions decoupled from population density? The approach used is based on the grid-cell level relationship between the Vulcan FF product and Landscan population (as shown in Figure 8a). However, despite this relationship (and its associated error statistics) only being valid for the United States and for 2002, it is used globally and for all years. Vulcan was used because it "relies least heavily on population", but justification for the large degree of temporal and spatial extrapolation is not provided. Moreover, while we can be fairly certain that modeling FF distribution purely on population will have spatial biases, we don't know that Vulcan is superior to other products such as ODIAC and EDGAR (versions 4.0

and later) that attempt to decouple population from FF by excluding large point source emissions. Comparing these products to gridded CDIAC would have been beneficial, because unlike Vulcan, they have global and interannual coverage. This approach would have specifically tested the uncertainties associated with not decoupling large point sources from the population proxy, a valuable contribution in my estimation. Overall, given the choice to use Vulcan, quantification of the impacts of using US-based error statistics for all countries is not discussed.

Clarifying language was added to section 4.3. Justification for spatial extrapolation is already stated in the text, Vulcan was used as it spatially covers the largest area and is least dependent on population for its carbon distribution. Justification for temporal extrapolation is more problematic, but Vulcan was again used as it is similar to most other FFCO2 maps in that a base year was done with most care and other years are mainly extrapolations from that base year. No better alternative was identified for temporal extrapolation by the authors or the reviewer.

We agree with the reviewer that the relative quality/fidelity of VULCAN versus ODIAC versus EDGAR is unknown. That is part of the reason for doing a study as reported in this manuscript. Spatially explicit uncertainty analyses will facilitate better comparisons of these and the CDIAC FFCO2 maps.

ODIAC and EDGAR were not used for comparison for different reasons. ODIAC is heavily derived from the CDIAC efforts and thus incorporates a heavy population distribution bias which is absent from VULCAN. EDGAR uses multiple proxies to distribute its FFCO2 emissions, many tied to population, and thus likely has a higher population distribution bias than VULCAN. Of course there are tradeoffs (especially in spatial and temporal coverage), but the authors decided to minimize the population bias since that is the effect trying to be isolated in section 4.3.

The ACP-destined text has been modified.

Because of this and the confusing explanation of the methodology, I can not recommend this manuscript for publication in its current form.

The authors obviously disagree with this recommendation or they would not have submitted the manuscript for publication on ACPD. The ACP-destined text has been modified to improve the clarity of the presentation which is the largest objection of the reviewer.

Other general points: a) Percent uncertainties are a bit misleading, because high percentages (e.g. at borders) may correlate with low absolute emissions. Looking at the range of % uncertainties can also be misleading. Looking at weighted means and standard deviations would be more meaningful. (e.g. in Table 4). b) There is a lot of description/discussion of `geography' uncertainty (4.2) even though this is a rather small component of total uncertainty, as acknowledged later. This section could be greatly condensed by, in particular, eliminating some of the examples of causes that uncertainty that certainly exist but are in fact very small contributors to total uncertainty - e.g. Colorado surveying errors, India/China border dispute (a region with low pop. density), and the northern protrusion of Minnesota into Canada.

a) The reply to the L547 comment below addresses this comment and is not repeated here. Table 4 already gives standard deviations. Weighted means might be another way to look at the uncertainties, if the weight were better defined by the reviewer. We give a weighted mean via the average in Table 4 where the weight is equal treatment to all populated cells which emit FFCO2. The manuscript has not been modified in response to this comment.

b) The reply to the L285 comment below addresses this comment and is not repeated here. The ACP-destined text has been modified.

c) As pointed out below, at several points in the manuscript other well known fossil fuel emissions products such as FFDAS, ODIAC and EDGAR are alluded to but not referenced explicitly. Being explicit would help the reader follow the arguments better.

We do name these, but regrettably not as soon as the reviewer would like. An earlier listing of the other products would interrupt the flow of those paragraphs and idea development in the manuscript. The ACP-destined text has been modified.

Line by line comments: L28. Is temporal uncertainty folded into the total uncertainty? If not, why state it in the abstract?

Yes, temporal uncertainty is folded into the total uncertainty. We account for temporal uncertainty at annual and monthly time scales for the respective data products. Temporal uncertainty is not separately evaluated in this manuscript, but is included in each of the three components: 1) tabular FFCO2 data (temporal uncertainty not discussed in this manuscript, but is partially accounted for in national uncertainty estimates), 1) annually varying geography maps, 2) annually varying population maps (when available) and negligible uncertainty introduced by the use of static population maps for the time periods considered. The manuscript has not been modified in response to this comment.

L42. What is the source of the statement that FF was the least uncertain component in the global carbon cycle? Please provide a reference.

Reference added to the ACP-destined text.

L53. Use of `significant' is vague here.

The reviewer is correct that significant is not explicitly defined here or previously in the manuscript. However, this is only the second paragraph of the introduction. The first paragraph is centered on global spatial scales. The second paragraph introduces temporal scales and brings the discussion to shorter magnitude spatial and temporal scales in order to better isolate specific components of the carbon cycle. “Significant” is intentionally vague in its dimensions, but it is clear that the dimension discussion is larger than point samples in both space and time. The manuscript has not been modified in response to this comment.

L54. By `direct samples' do you mean `air samples'?

Direct samples could be air samples, but the term is much broader than that and includes leaf clippings, water samples, ... Direct samples are any samples of the global carbon cycle. They are limited in time and/or space. Thus, while they may contain valuable and useful information, that information is usually better understood in broader spatial and temporal contexts. Hence, models are used to determine histories of the samples that are consistent with sample properties. The manuscript has not been modified in response to this comment.

L65. It would be useful to define `component fluxes' clearly here. I presume this means land, ocean, fossil and land use-change fluxes, but I'm not sure.

The ACP-destined text has been modified.

L66. Please provide references for the statement `present efforts in gridding ....'

This sentence is general in nature and thus no reference is given here. However, five other gridded products are specifically named and referenced in the next paragraph. The L66 sentence certainly applies to these five data products. Listing these five other gridded products in L66 would be redundant. The manuscript has not been modified in response to this comment.

L70. Change to `advances'.

The ACP-destined text has been modified.

L75. To eliminate confusion with the idea of active sequestration of fossil fuel CO2, I would simply say `carbon flux from the fossil fuel reservoir'.

The ACP-destined text has been modified.

L92. This is an opportunity to cite some other gridded FF products.

Five examples of other gridded FF products were just referenced in the preceding paragraph. The proxies listed around L92 are sourced from those products. Relisting those products here would be redundant. The manuscript has not been modified in response to this comment.

L95. For consistency with earlier language, change `stock' to `reservoir'.

The ACP-destined text has been modified in repsonse to this comment with several changes throughout this paragraph.

L95-L101. This paragraph is not necessary and interrupts the flow of the discussion. Moreover, it should be noted that significant seasonal and spatial gaps in satellite CO2 products preclude their use to quantify the atmospheric carbon reservoir.

The authors believe this paragraph is important to help define the scope of this manuscript. Satellite measurements have often been suggested as a solution to quantifying FFCO2 emissions. The reviewer mentions one such deficiency with the current satellites. Another deficiency includes present satellites do not directly measure FFCO2, but total CO_2 in the atmosphere; models are then needed to tease out the FFCO2 signal from the total CO_2 signal. Other deficiencies also exist with the current satellites. These deficiencies are not discussed in this manuscript paragraph as they are not central to the focus of this manuscript. The manuscript has not been modified in response to this comment.

L115. `Stable carbon maps' would actually include both 13C and 12C. Replace with something like "maps of the 13C:12C ratio (δ?13C) of CO2 ...'

The ACP-destined text has been modified.

L150. Strike `exact' - redundant.

The ACP-destined text has been modified here and in Fig. 1.

L175. Add the word `degree' or use the degree symbol after 1x1.

1x1 is defined (including the word "degree") in the line immediately above this line. 1x1 is also used in many places further in the manuscript. The manuscript has not been modified in response to this comment.

L177. "geography data" and later "geography map" are confusing terms because `geography' can refer to landforms, population, etc. What is meant here is `political [or national] boundaries', and I suggest using something along those lines.

The first two uses of the word "geography" in the manuscript clearly link the word to national borders. L177 is the first instance where this link is not clearly established within the same sentence, but "geography" is clearly linked with "attributing grid cells to a single country " two sentences before L177. While it is true that geography is a multi-faceted science that includes historical (e.g., political borders), physical (e.g., landforms), and cultural (e.g., population) subfields, the manuscript clearly identifies that geography map is referring to a national borders map. The first paragraph of section 4.2 makes this association explicitly clear. The manuscript does not use the word geography in its physical or cultural contexts. The rest of the manuscript was searched for the word "geography" to ensure that there was no ambiguity in the association with political units where it was implied. The manuscript was modified in one instance in response to this search.

L196. Are there no population density maps for specific years (or even, say, five year periods) prior to 1989?

As stated in the manuscript, the authors are not aware of a "better alternative". The manuscript has not been modified in response to this comment.

L200. At a 1x1 resolution, how different are nighttime and daytime populations? Does this change have any impact?

There are differences between nighttime and daytime population distributions. While the global total for each data product differ by more than 100 million people, it is not the absolute population that is of concern here, it is the relative population distribution within each country that is of concern here. The most obvious differences are where one population data product has a positive value where the other has a zero value. Comparing calendar year 2000, the only year for which both have original data (not interpolated data), there are also differences in relative population values. Most populated grid cells are less than 10% different between daytime and nighttime relative populations. The ACP-destined text has been modified.

L210. For clarity, add `fixed' between `CDIAC' and `population'.

Those words do not appear on L210, but do on 216. L216 of the ACP-destined text has been modified.

L214. Delete `degree'.

Reading the text now does make this phrasing (i.e., second degree) seem odd. However deleting the word loses the original intent of conveying seconds is used as a distance unit here, not a time unit. The ACP-destined text has been modified to preserve this original intent.

L219. Change `present' to `2011', the last year of the present study, i.e. opposed to 2016 or later.

This comment has merit, but the authors intentionally used present so as not to bound (in time) or date the manuscript. Earlier in the manuscript, present is defined in terms of the latest year of the data set as calendar year 2011 (i.e., the first sentence of the seventh paragraph of the Introduction). When writing this manuscript, the authors were actively working on a newer release of the gridded data sets and associated uncertainty maps (as was stated in the "Data availability" section.). That work was completed and these maps now go through calendar year 2013. Putting in the year 2011 or 2013 would unnecessarily bound the date in the manuscript. The manuscript has been edited in many places to update the 2011 to 2013. The Data availability section has also been edited to reflect web-availability of the data as was promised in the ACPD version of this manuscript. The ACP-destined text has been modified.

L227. Delete `exact'.

The ACP-destined text has been modified.

L234. The distinction between `specific' and `general' uncertainty is unclear.

"Specific" is used here to denote the uncertainties in the specific data used. For example, the population used as a proxy for FFCO2 distribution (e.g., Fig. 8). "General" is used here to denote the uncertainties in the functional role that the input data provide. For example, while we use population as an FFCO2 proxy, others may use power plants as an FFCO2 proxy. The specific uncertainty would be a power plant-C relationship akin to what is seen in Fig. 8. The general uncertainty would be that the use of a power plant proxy contributes uncertainty at all. At present, most users of power plant proxies do not consider locational, temporal, or FFCO2 magnitude errors occur in that proxy - they are glad to have a feature-specific proxy to use at all. The authors have provided the specific uncertainties for the three inputs (e.g., Figs. 3, 6, and 8) used here. The authors have also provided text throughout section 4 on the general uncertainties (including, but not limited to production versus apparent consumption data in section 4.1, different spatial representations in section 4.2, newer data sets/proxies in section 4.3). The manuscript has not been modified in response to this comment.

L264. Is the discussion of monthly emissions estimates and their uncertainty relevant to this paper? It's not clear to me reading the paper whether the uncertainty analysis pertains to annual or monthly global FF maps. On the one hand this paragraph discusses monthly uncertainties, but at the end of this section (4.1) it is clearly stated that "There are 62 uncertainty assessments completed for the 1950-2011 time series ..." (i.e. not 62x12 uncertainty maps.)

Yes, monthly temporal intervals are also relevant to this manuscript (in addition to annual temporal intervals). CDIAC produces both annul and monthly FFCO2 mass emission maps. This manuscript describes the 1x1 uncertainty maps associated with the annual AND monthly 1x1 mass emission maps. As there is a lot of overlap in their formulation, one manuscript suffices instead of writing two similar manuscripts - one for annual maps and one for monthly maps. The first use of the word "annual" in this manuscript occurs in the Introduction and is followed two sentences later with the first use of the word "monthly". The next sentence states "With the global FFCO2 emission uncertainty analysis completed, a gridded uncertainty analysis can be applied to the annual and monthly maps " clearly indicating that the reader should expect two temporal intervals to be discussed later in the manuscript.

Annual and monthly temporal intervals are addressed in each of the three component uncertainty sections: 4.1, 4.2, and 4.3. However, most of the text in the manuscript is directed toward the annual temporal interval as this the most common use of the CDIAC data. Where significant deviations occur for monthly data, more manuscript text is devoted (e.g., section 4.1). Most of the manuscript text applies to both annual and monthly temporal intervals.

The ACP-destined text has been modified around L264 to clarify annual and monthly uncertainty maps exist.

L285. As mentioned above, condensing this section (currently ~100 lines) would improve the flow of the paper, allowing the reader to get to the most relevant methods and findings more quickly.

The authors believe that the 12 paragraphs in this section are important. They cover: introduction, raster maps, raster uncertainty, political issues, rescaling, uncertainty contribution, uncertainty algorithm, limitations, 100% uncertainty, intermediate uncertainty, results, and general uncertainty expansion. All of these contribute ideas central to understanding the role geography plays in creating uncertainty in mass emission maps, both in the CDIAC case and for similar data products.

Text in the Fig. 4 captions was repetitive of text in section 4.2; three lines of text were deleted. The main body of section 4.2 was reviewed and not modified in response to this comment. The

ACP-destined text has been modified.

L393. Change to `assumes each country has'

The ACP-destined text has been modified.

L397. References to/descriptions of the FF products that use these approaches would be useful.

Reference examples have been provided. Descriptions have not been provided as this is a general list. The ACP-destined text has been modified.

L408. The second sentence of this paragraph seems too obvious to be worth including.

It may seem obvious to some, but not to others. The first author has spoken similar words at meetings in response to questions when presenting versions of this work. So the answer is provided here for future similar questions. The manuscript has not been modified in response to this comment.

L417. As pointed out above, this assessment is only carried out (and valid) for the United States.

Comment well taken. We can discuss the validity of the population-carbon relationship in the future. However, this is the best population-carbon data set the authors could identify and thus they used it. It is far from perfect. Text has been added to the ACP-destined manuscript to emphasize the U.S. to global extrapolation at the end of the paragraph, two paragraphs ahead in the manuscript (i.e., The lower panel of Fig. 8...).

L431. Strike `axes'.

The ACP-destined text has been modified.

L431. Because of the very large y-axis range in Fig. 8b, it's very hard to see how much non-linearity there is.

Figs. 8b and 8c have had their y-axes redrawn to enlarge the physical space they use by expressing population in millions. But this may not fully address the concern of the reviewer. However, the authors believe the non-linearity is obvious. If the data were linearly related, the data would fall on either of the two linear fit lines in Fig. 8b. The points do not and the accompanying confidence intervals are large. The manuscript has not been modified in response to this comment.

L432. How was the best fit line and its 2-sigma confidence intervals calculated? In addition to using the values generated by a software package, one could use a bootstrap or other techniques; different approaches may be more or less sensitive to outliers or have other beneficial characteristics. Additionally, there may be an effect on the slope (assuming a "model-I" regression was used) by plotting FFCO2, the more uncertain value, on the y-axis compared to the x-axis. Was this effect investigated by switching the axes?

The sequence of events for data analysis was as follows:
1. Original data (linear scale) was converted to ln scale (Fig. 8a) as it is not obvious that a linear fit is an appropriate model to use when the data are portrayed on linear scales.
2. Linear fit and 95% confidence intervals were calculated with the TableCurve 2D v 5.01 software package. Other packages were also tested, but the same linear fit was obtained. Outliers were included in this fitting procedure.
3. Original data and three curves of step 2 (i.e., linear fit and 95% confidence intervals) were plotted on linear scales in Fig. 8b.
4. Monte Carlo simulations were performed to provide constrained linear fit and 95% confidence intervals. MC simulations were executed to maintain the total carbon of the system. A robust estimate of the data distribution was obtained by excluding original data outliers. The result of the MC process is the same general tendency of the FFCO2-population relationship (as evidenced by the very similar linear fit lines), but with a smaller standard deviation associated with the constrained fit line (obtained by the exclusion of original data outliers).
5. Fig. 8c was constructed from Fig. 8b constrained fit results.

The authors do not understand the reviewer's switching of axes comment. FFCO2 was the y-axis variable in the regression analysis and population was the x-axis variable. The authors did not investigate switching the axes of the variables for a regression analysis because that is not consistent with the mapping procedure. The mapping procedure uses population as the predictor for FFCO2 emissions. The procedure does not use FFCO2 emissions as a predictor of population which is what the switched axes regression analysis would test.

The manuscript has not been modified in response to this comment, but this line was modified in response to the second L436 comment.

L435. The description of the Monte Carlo process to produce a "constrained" 2-sigma confidence interval is confusing to me.

See the "sequence of events" in the comment immediately above. Since this is essentially equivalent to the text already in the manuscript, the manuscript has not been modified in response to this comment.

First (L436), why does the selection of a population need to be random? Couldn't all the grid cell population values in Fig 8 (a or b) be dealt with sequentially?

The authors tested both population-selection methods the reviewer mentions. The 1000 MC results come out virtually the same in each case. The sequential access runs more quickly than the random access as the constant carbon requirement is met with fewer runs (i.e., the random access method results in more runs that do not meet the constant carbon requirement and thus more runs are needed to meet the 1000 successful run requirement; this is likely the result of the specific algorithm programmed). Only about 6% of the MC runs result in the constant carbon requirement when run in sequential access mode, so to get 1000 successful MC runs requires approximately 16000 MC attempts. The ACP-destined text has been modified.

Second, the statement `selecting an adjustment...in accordance with a robust 2 sigma interval...'

is unclear. A) What is meant by `robust'? Because of this, it is also hard to understand why `the robust 2 sigma interval minimizes the effect of outliers.' B) The method of the adjustments to the `regression fit FFCO2' is also not totally clear. Are these adjustments random draws from the 2 sigma intervals calculated in the log-log fit shown in 8a? (And is this what is repeated 1000 times?) If the 2 sigma intervals were calculated from a linear fit, this process would be straightforward, but to make adjustments to FFCO2 (not ln(FFCO2)), the original 2 sigma distribution needs to be transformed, which is not straightforward. Is it that both the random draws and the adjustments take place in log-log space, and then are transformed to FFCO2 space?

A) "robust" is defined in the next sentence as it pertains to reducing the effect of outliers. To reduce the effect outliers, the constrained linear fit was determined with the original data points that fell within the +/- 2 sigma interval of the original data series. This excluded 56 original data points (kept 793 data points) in the determination of the constrained linear fit. The standard deviation later used in the MC adjustments comes from this 793 data point set. Other data exclusion formulas were tested, but this one was selected for final use due to its simplicity and end results as seen in Fig. 8c.

B) For all 849 populations, using the linear fit equation of the 793 point data set, an FFCO2 value was calculated. This value was then adjusted by the standard deviation calculated from the 793 data point set (see A) immediately above) multiplied by a randomly selected standard deviation interval ranging from -3 to +3 sigma in accordance with the distribution of a normal curve (i.e., ~68.2% of s.d. intervals fall within +/- 1 sigma, ~95.4% of s.d. intervals fall within +/- 2 sigma, and ~99.7% of s.d. intervals fall within +/- 3 sigma). The adjusted value is kept as a pair with the population. After the 1000 MC runs are met (with all runs meeting the constant carbon criterion), the average and standard deviation FFCO2 emission at each of the 849 populations is determined. To be clear, the average and standard deviation are determined with a sample size of 1000 at each population.

The constrained linear fit calculation and adjustment can be done in ln space or linear space, the two being related by ln and exp functions. The results are the same regardless of which space is used, within the accuracy of the ln and exp transformations. For this analyses, linear space was used as Fig. 8c presents the final results in linear space (which is the space more familiar to most people working with populations).

Ln space is only used for Fig 8a of this analysis as it is the beginning for the use of a linear fit to the data. Linear space is used for the rest of the data analysis.

The ACP-destined text has been modified.

Third, I do not understand the mechanism by which the sum of FFCO2 from all grid cells is made to equal the national total. After calculating an FFCO2 distribution associated with the population of each 1x1 grid cell, is there a selection process to ensure that the total FFCO2 of all cells equals the national total (and thereby creating the constrained fit)?

Comment addressed above. The ACP-destined text has been modified.

L442. Given that there is a distribution of 1000 elements, wouldn't it be more straightforward to calculate the 2.5 and 97.5 percentile bounds, to allow for asymmetry? (or is this what is meant by the present use of `sigma'? If so, it would be clearer to refer to confidence intervals instead of sigmas, which carry the connotation of symmetry.).

As stated in the third to last paragraph of the introduction, the authors assume the uncertainty is symmetrical about the central estimate. The connotation of symmetry is intended. The manuscript has not been modified in response to this comment.

L445. The 2 sigma intervals in the upper panel appear symmetrical about the best-fit line, contrary to what is stated.

They may appear symmetrical, but they are not. The two curves have different average slopes (and the slopes on one curve are not constant). The manuscript has not been modified in response to this comment.

L446. What is the break point between low and high population values in determining whether to use the +2 sigma or -2 sigma curves.

The breakpoint is simply the result of a distance measurement. It occurs at a population of about 40,000 people. Since this exact value has no lasting impacts on the manuscript, it has not been reported in the manuscript. The manuscript has not been modified in response to this comment.

L475. Is there any expectation of negative populations or emissions?

The authors do not expect negative populations. People in this field of study do discuss negative emissions (as a sink, usually in response to a carbon capture technology implementation). These two observations are given as basic characteristics of the plot and data set. The manuscript has not been modified in response to this comment.

L477. `Fourth...' How could this not be true if there are no negative populations of FF emissions - both of which should be obvious.

Not obvious to all. Some maps have FFCO2 emissions in regions of zero population. Zero is neither negative or positive. Again, this is an observation or basic characteristic of the data set. The manuscript has not been modified in response to this comment.

L478. `Fifth...' How is this effect shown in Fig. 8b?

Agreed, the effect is not shown in Fig. 8b. The ACP-destined text has been modified.

L487. Change to `These maps' to be consistent with previous sentence.

The ACP-destined text has been modified.

L488. Instead of assigning uncertainty to zero, could single grid cell countries be assigned uncertainties linked to their population (via Fig. 8c)?

Single grid cell countries have no ambiguity introduced into their FFCO2 emission distribution as a result of their population since their population only occupies one grid cell. There may be ambiguity as their population may be in the wrong grid cell, but that is addressed by the geography map uncertainty. The manuscript has not been modified in response to this comment.

L501. Recommend inserting `that' after `was'.

The ACP-destined text has been modified.

L506. Change `this' to `which'

The ACP-destined text has been modified.

L533. Could it also be that the 2011 population maps have higher resolution than the 1950 ones? For example, in the 1950 map, Australian states appear to have uniform population density, as do some countries like India.

The spatial resolution of both the 1950 and 2011 population maps is the same: 1x1. The appearance of the 1950 map having more homogeneous populations is a function of the scale used to portray the data (e.g., gradations of 20%). Different scales will give different appearances. The manuscript has not been modified in response to this comment.

L547. I'm not sure of the value of the breakdown of what fraction of cells have uncertainty dominated by which uncertainty component, because in many cases these cells may have very low emissions.

The value of the breakdown is to give future research guidance on which components of the uncertainty are in need of the most urgent work. The population proxy contributes the most uncertainty to the most number of cells. This issue is addressed in the Discussion section of the manuscript; this begins two paragraphs later in the manuscript. The value of the breakdown to current readers of the manuscript is to inform them that all three components contribute uncertainty and that to focus solely on one component (e.g., the population proxy) gives an incomplete understanding.

The reviewer is correct that the uncertainty maps are in percentage units, not absolute mass units. The authors leave it to the readers to calculate absolute mass units by multiplying the FFCO2 mass map by the corresponding FFCO2 uncertainty map. The authors did not think it necessary to include an absolute mass uncertainty map in the manuscript. Here, the reviewer focuses on low emission cells. Other reviewers or readers might focus on high emission cells (e.g., large cities) or their geographic region of interest. To fully understand these relationships, one needs the FFCO2 mass emission map, the FFCO2 percentage uncertainty map, and the FFCO2 absolute unit uncertainty map. The data accompanying the manuscript allows the reader to obtain all three maps.

The manuscript has not been modified in response to this comment.

L558. Given annual maps, isn't the time scale of this example much to fast too matter?

The authors discuss both annual and monthly maps in the manuscript. Others are discussing/creating weekly, daily, and finer time scale maps. Power plant shutdowns can last hours to days to months to more than a year. The time scale of this example is applicable to the maps presented in this manuscript and to other maps available in the broader community. The manuscript has not been modified in response to this comment.

L585. Replace `Other map distributions...' with specific examples like ODIAC, FFDAS, EDGAR, etc.

Examples of other maps are named earlier in the manuscript. It is not necessary to make a partial call to these maps again. This comment is more general in nature, and is not meant to call out one specific or several specific maps. The manuscript has not been modified in response to this comment.

L609. Which efforts has CDIAC supported? References?

CDIAC has supported VULCAN, ODIAC, FFDAS, and others through direct funding for travel and consultations. The motivation for this support from CDIAC is to obtain the best estimates of FFCO2 distribution, this includes investigation of alternative distribution algorithms. These details are not provided in the manuscript as they do not add to the development of the ideas in the manuscript. Evidence of this support comes through coauthorships (e.g., Hogue et al., 2016 cited in this manuscript) and acknowledgments in the publications of these other efforts. The manuscript has not been modified in response to this comment.

L612. Change `a honest' to `an honest'

The ACP-destined text has been modified.

L615. Provide references for the 36% value and for the subsequent 15-20% value.

The ACP-destined text has been modified.

L626. Regarding INFLUX, would comparison to results of Turnbull et al, JGR 2014 be better since they were able to focus just on FFCO2?

The first author spoke to Turnbull about this issue. Her reply was that her previously published work does not cover the entire city and sent us to the Cambaliza at al. paper which we cite. The manuscript has not been modified in response to this comment.

L649. Please provide a reference for the `known spatial deficiencies'.

The known spatial deficiencies are not a criticism of the CARMA database, locations in

CARMA were determined by postal code and that was fine for CARMA purposes. The problem arises when CARMA is repurposed for other uses such as emissions mapping. Then, locational issues become more prominent because postal codes are not always co-located with stack locations. Postal code can refer to stack location, plant entrance, plant headquarters, ... Gurney, Oda, Marland, and others have all verbally discussed the locational issues with CARMA. Three references have been supplied, but much more information exists verbally and in meeting presentations. The ACP-destined text has been modified.

L666. How were the population and political border uncertainties reduced? Arbitrarily?

As explained later in the paragraph, population uncertainties were reduced to the minimum line of Fig. 8 and geography uncertainties were reduced to 50% of Fig. 6 values. Lacking further information, the minimum line of Fig. 8 may be more appropriate than the average line used previously in the manuscript. Fifty percent is roughly intermediate between 100% used previously in the manuscript and the 10% of Hogue et al. (2016). These were not arbitrary decisions. The manuscript has not been modified in response to this comment.

L669. Why does one tenth of a grid cell correspond with 10% uncertainty?

Since the grid cells used here have a uniform population density within each cell, moving the borders of a cell 10% spatially results in approximately ~10% uncertainty. This is approximate because the exact uncertainty is dependent upon the bordering cells population value also. The ACP-destined text has been modified.

L685. Is it fair to compare uncertainties for 0.25 deg. Vs. 1.0 degrees? Using a higher resolution population database along with Vulcan could result in higher uncertainties in this study and an even bigger contrast with Rayner et al.

A valid question and one without an easy answer. As the reviewer points out, the work presented here and that of Rayner et al. have different spatial resolutions (as the text explicitly mentions). However, the comparison remains for two reasons. First, there is close agreement between the two works in absolute magnitude. Second, differences between the two uncertainty analyses are specifically named. As the reviewer mentions, putting the two works on the same basis (e.g., spatial resolution, population uncertainty treatment, clipping, number of Monte Carlo runs, ...) may result in a better or worse absolute magnitude uncertainty comparison. The calculations necessary to put the two works on the same basis have not been performed in response to this comment as 1) it is outside the scope of this manuscript, 2) if done, the results are not representative of an actual data product available, and 3) the component data of the non-CDIAC data product are not freely available. One could delete this entire paragraph, but the comparison (for the two reasons described above) justify keeping the paragraph. The manuscript has not been modified in response to this comment.

L735. Decrease in uncertainty with aggregation could be tested directly using the methods in this paper by simply aggregating both Vulcan and Landscan to, 2x2, 3x3 degrees, etc.

It is not as straightforward as the reviewer suggests. Figure 12 shows total uncertainty and the test the reviewer suggests modifies only one component of that total uncertainty. A more complete treatment would also entail modifying the geography map so that it agrees with the population map. This would then entail modifying the carbon distribution algorithm for handling the new spatial resolutions. The new spatial resolutions would be particularly difficult for small countries. The calculations necessary to test the comment of the reviewer have not been performed as 1) it is outside the scope of this manuscript, 2) if done, the results are not representative of an actual data product available, and 3) many decisions would be necessary along the calculation pathway that would need to be justified and explained for the calculation results to have merit. Different decisions could lead to very different results. The manuscript has not been modified in response to this comment.